

EXTERNAL SANDHI IN PĀLI (WITH SPECIAL REFERENCE TO THE *SUTTANIPĀṬA*)

Since words in Pāli can, as a general rule, end only in a vowel or anusvāra, external sandhi in Pāli consists of a series of contractions of final vowels with the initial vowels of following words, or the loss of the nasalisation and the contraction of the vowel remaining.

As Geiger points out,¹ external sandhi in Pāli differs fundamentally from that in Sanskrit. It is always arbitrary and applies only to words which are syntactically closely connected. It permits hiatus, elides and combines vowels including nasalised vowels, and in particular depends upon the crasis or elision of *-i* and *-u* before vowels rather than the development to *-y* and *-v* found in Sanskrit. Windisch pointed out that such a sandhi system appears to be older and more natural than the system found in Sanskrit.²

When the final vowel is *-i/-e* or *-u/-o* we do sometimes find *-y* or *-v* appearing. Since, in the case of *-e* and *-o*, this is not historically correct, I regard these as hyperforms, with *-y* and *-v* inserted incorrectly by someone who had some knowledge of Sanskrit. I similarly regard the appearance of *-y* and *-v* in place of *-i* and *-u* as insertions in imitation of Sanskrit.

On occasion, the long vowel which results from combination remains before a doubled consonant or consonant group (indicated by *CC* in the following examples), in defiance of the Law of Two Morae, by which a long vowel should be shortened in those circumstances. There seems to

¹ W. Geiger, *Pāli Literatur und Sprache*, 1916, § 68.

² Quoted by Geiger, *ibid*.

be no example in the Suttanipāṭa of this happening before a nasal + consonant (indicated by *NC*).

The great variety of sandhi forms which occur probably supports Windisch's belief. External sandhi in Pāli is, however, by its very variety more complex and confusing to readers, since it is not always obvious what vowels are being combined. In the belief that readers of Pāli, especially beginners,¹ might find an analysis of sandhi forms helpful, I have collected here a number of examples of external sandhi which occur in the Suttanipāṭa, although I have very occasionally given examples from other texts of sandhi forms which do not occur in the Suttanipāṭa. I make no claims for completeness, but I believe that the list contains most of the sandhi forms which readers of Pāli are likely to encounter. For the most part I give only one example of each sandhi development, usually but not always the first occurrence of the form in the Suttanipāṭa. Numbers without any other indication refer to the verses of the Suttanipāṭa, and references prefixed with p. refer to the prose portions of that text. I sometimes give additional examples in the footnotes, including references to texts other than the Suttanipāṭa, where the sandhi is rare, or for some reason particularly interesting.

Since editors sometimes differ in the way in which they show contractions and elisions, I have omitted inverted commas, etc., and simply show the result of sandhi between hyphens. Since we cannot always be certain what vowels have been elided, I have drawn attention to possible ambiguities. As a result of the same uncertainty, some of my reconstructions may be wrong. I hope that the number of incorrect forms has been kept to a minimum.

¹ Some of whom, I know, find the detailed accounts of sandhi in Geiger, §§ 68 foll. and A.K. Warder, *Introduction to Pali*, pp. 213 foll., rather daunting.

The developments of the particles *api* and *iti* sometimes show a secondary development from the historical Sanskrit sandhi. I have discussed these in a separate excursus.

In an Appendix at the end I list the sandhi developments in reverse order, so that beginners, faced with a sandhi form which they do not recognise, may have guidance in recognising which vowels may have combined to produce the form in question.

-a)	-a + a- > -ā-	14 yassa + anusayā > yassānusayā
	-a + aCC- > -aCC-	p. 16,6 yassa + atthāya > yassatthāya
	-a + aCC- > -āCC-	5 na + ajjhagamā > nājjhagamā ¹
	-a + aNC- > -aNC-	6 yassa + antarato > yassantarato
	-a + ā- > -ā-	p. 16,5 pana + āyasmā > panāyasmā
	-a + i- > -ā-	84 Cunda + iti > Cundāti ²
	-a + i- > -i-	p. 106,12 tassa imāni > tassimāni
	-a + i- > -ī-	793 kena + idha > kenīdha
	-a + i- > -e-	272 nigrodhassa + iva > -asseva ³
	-a + iCC- > -iCC-	39 yena + icchakaṃ > yenicchakaṃ
	-a + iNC- > -iNC-	516 yassa + indriyāni > yassindriyāni
	-a + u- > -u-	p. 13,5 tena + upasaṃkami > tenupa-
	-a + u- > -ū-	433 na + upasussaye > nūpasussaye ⁴
	-a + u- > -o-	221 na + upeti > nopeti ⁵
	-a + uCC- > -uCC-	435 pattassa + uttamavedanaṃ > pattassuttama-
	-a + e- > -e-	81 na + esa > nesa

¹ cf. 8 foll. na + accasārī > nāccasārī; 291 na + assu > nāssu.

² This is probably not the result of -a + i-, but is an analogical form based upon -i + iti. See the note on -i + i- below.

³ cf. 768 sappassa + iva > sappasseva.

⁴ cf. 801 yassa ubhayante > yassūbhayante; 849 na + upasaṃkheyyo > nūpa-; 1077 jivītena + upapannaṃ > jivītenūpa-; 1106 ca + ubhayaṃ > cūbhayaṃ.

⁵ cf. 779 na + upalitto > nopalitto; 812 na + upalippati > nopalippati.

-a + o- > -o-	270 iva + ossajanti > ivossajanti ¹
-am) -am + a- > -ā-	p. 16,1 labheyyam + aham > labheyyāham ²
-am + a- > -a-	538 -tamam + agā > -tamagā
-am + aCC- > -aCC-	225 samam + atthi > samatthi ³
-am + aCC- > -āCC-	769 gavam + assam > gavāssam ⁴
-am + aNC- > -aNC-	1148 pañhānam antakaro > pañhānantakaro
-am + ā- > -ā-	285 brāhmaṇānam + āsum > brāhmaṇānāsum
-am + i- > -ī-	833 yesam + idha > yesīdha
-am + u- > -u-	700 anagāriyam + upetassa > anagāriyupetassa
-am + u- > -ū-	901 tapam upanissāya > apūpanissāya ⁵
-am + uCC- > -uCC-	248 niccam + uyyutā > niccuyyutā
-ā) -ā + a- > -ā-	53 yathā + abhirantam > yathābhi-
-ā + aCC- > -aCC-	146 pāṇabhūtā (perhaps) + atthi > -bhūtatti ⁶

¹ Is this iva + osajanti (with gemination of -s-) or iva + ussajanti, i.e. is it -a + o- > -o-, or -a + u- > -o-?

² cf. 353 1061 vācam abhikaṅkhāmi > vācābhikaṅkhāmi; 365 nibbānapadam + abhipatthayāno > -padābhipatthayāno; 788 etam + abhijānam > etābhijānam; 896 khema + abhipassam > khemābhipassam; 1033 jappam + abhilepanam > jappābhilepanam; 1057 1083 etam abhinandāmi > etābhinandāmi.

³ cf. 693 vitthārikam + assa > vitthārikassa; 970 kuvam + ajja > kuvajja [NB kuvam in 970a].

⁴ Is this internal or external sandhi, i.e. is it a compound or two separate words? Is it -a + aCC- > -āCC- or -am + aCC- > -āCC-?

⁵ cf. 972 kukkuciyam + upacchinde > kukkuciyūpa-.

⁶ cf. 151 yāvatā + assa > yāvatassa; 508 Brahmā + ajja > Brahmajja; 595 padakā + asmā > padakasmā; 875-76 ettāvatā + aggam > ettāvataggam; 1022 unṇā + assa > unṇassa.

-ā + aCC- > -āCC-	379 mantayitvā + ajjhagamā > mantayitvājjhagamā ¹
-ā + aNC- > -aNC-	570 dantā + amha > dantamha
-ā + ā- > -ā-	249 mantā + āhutī > mantāhutī
-ā + i- > -ī-	182 saddhā + idha > saddhīdha
-ā + iCC- > -iCC-	512 manasā + icchasi > manasicchasi ²
-ā + iNC- > -iNC-	229 yathā + indakhīlo > yathinda-
-ā + e- > -e-	19 Mahiyā + eka- > Mahiye-ka-
-ā + o- > -o-	60 yathā + odhikāni > yathodhikāni
-i) -i + a- > -a-	90 hi + apeti > hapeti ³
-i + a- > -ā-	p. 15,1 kassaci + aham > kassacāham ⁴
-i + a- > -i-	783 iti + ahan > itihan
-i + a- > -tyā-	Ja VI 325,26* ti + abhivadī > tyābhi-
-i + aCC- > -aCC-	672 api ⁵ + atthi > apatthi
-i + aCC- > -iCC-	162 kacci + assa > kaccissa
-i + ā- > -ā-	287 sayanehi + āvasathehi > sayanehāvasathehi
-i + i- > -i-	381 keci + ime > kecime
-i + i- > -ī-	27 gavampati + idha > -patīdha ⁶

¹ cf. 830 sā + assa > sāssa; 1100 āsavā + assa > āsavāssa.

² cf. 1030 manasā + icchatha > manasicchatha.

³ cf. 231 catūhi + apāyehi > catūhapāyehi.

⁴ cf. 693 phusissati + ayam > phusissatāyam; 693 vattessati + ayam > vattessatāyam; and Thī 206 tehi + anucinnam > tehānucinnam.

⁵ See K.R. Norman, *The Group of Discourses* (= GD) II, note ad 672.

⁶ cf. p. 13,15 bhuñjāmi + iti > bhuñjāmīti. The development of vowels before (i)ti follows a consistent pattern, and I have listed them in their appropriate places (e.g. -a + i- > -ā-; -i + i- > -ī-; -u + i- > -ū-; -e + i- > -e-; -o + i- > -o-), but since the development of -am > -an in comparable situations (e.g. 9 idam + ti > idantī) seems to suggest that we are dealing with ti, not iti, I would rather suggest that, on the analogy of historical Skt -i + i- > -ī-, e.g. bhavati + iti > bhavātīti which was analysed as bhavātī ti, short vowels before (i)ti were lengthened, and -e and -o remained unchanged. A similar analogical development has taken place in

-i + e > -e-	52 <i>pi + etāni</i> > <i>petāni</i> ¹
-i + e > -ye-	Ja VI 267,27* <i>ti + eke</i> > <i>tyeke</i>
-iṃ)	
-iṃ + a- > -ā-	Vin IV 162,1 <i>kiṃ + ahaṃ</i> > <i>kāhaṃ</i>
-iṃ + aCC- > -iCC-	Ja V 362,25* <i>añjaliṃ + assa</i> > <i>añjalissa</i>
-iṃ + i- > -ī-	1078 <i>munīṃ + idha</i> > <i>munīdha</i>
-u)	
-u + a- > -ā-	482 <i>sādhū + ahaṃ</i> > <i>sādhāhaṃ</i> ²
-u + a- > -va-	Ap 516,14 <i>bhavesu + ahaṃ</i> > <i>bhavesvahaṃ</i> ³
-u + ā- > -ā-	p. 31,18 <i>sādhū + āvuso</i> > <i>sādhāvuso</i>
-u + ā- > -vā-	424 <i>kāmesu + ādīnavaṃ</i> > <i>kāmesvādīnavaṃ</i>
-u + i- > -ū-	173 <i>su + idha</i> > <i>sūdha</i> ⁴
-u + i- > -ī-	1082-83 <i>su + idha</i> > <i>sīdha</i>
-u + e- > -ve-	p. 104,13 <i>adhivāsetu + eva</i> > <i>adhivāsetveva</i> ⁵

Ardha-Māgadhī, where the historical *bhavaṭīti* was replaced by the alternative *bhavatitti*, analysed as *bhavati tti*, which led to the analogical -a *tti* and -u *tti*. Before the double consonant a long vowel is shortened, e.g. *samīti tti* (Uttarādhyāyanaśūtra 8.9), for *samitē tti*, *logu ti* (Utt. 8.20) for *logō tti*, and *patta tti* (Utt. 12.47) for *pattā tti*.

¹ cf. p. 124,5 *hi + evaṃ* > *hevaṃ*, where, however, *hevaṃ* may be a word in its own right (see K.R. Norman, "Pāli lexicographical studies IX", *JPTS* Vol. XVI, 1992, pp. 77-85 [p. 85]); p. 218,18 M III 133,29 *ti + eva* > *teva*.

² For 1032 *kissābhilepanaṃ* see *GD* II, note ad 1032.

³ I suspect that this is a Sanskritism for *bhaves' ahaṃ*.

⁴ p. 13,13 *bhuñjasūti* is probably not the result of -u + i-, but is an analogical form based upon -i + iti. See the note on -i + i- above.

⁵ I suspect that this is a Sanskritism for *adhivāset' eva*; cf. 611 *manussesu + etaṃ* > *manussesvetam* (= *manusses' etaṃ* ?); p. 126,12 *tu + eva* > *tveva*; and Thī 499 *pitupitusu + eva* > *pitupitusveva*, which is probably a Sanskritism for *pitupitus' eva*. See *GD* II, note ad 144.

-um)	-um + a- > -ā-	D I 213,10 <i>amum + ahaṃ</i> > <i>amāhaṃ</i> ¹
-e)	-e + a- > -ā-	891 <i>ye + abhivadanti</i> > <i>yābhivadanti</i> ²
	-e + a- > -e-	891 <i>te + abhirattā</i> > <i>tebhirattā</i>
	-e + aCC- > -aCC-	377 <i>te + atthi</i> > <i>tatthi</i>
	-e + aCC- > -yaCC-	340 <i>te + atthu</i> > <i>tyatthu</i>
	-e + aCC- > -yāCC-	961 <i>ke + assa</i> > <i>kyāssa</i>
	-e + aNC- > -aNC-	62 <i>salile + ambucārī</i> > <i>salilambu-</i>
	-e + i- > -e-	137 <i>me + idaṃ</i> > <i>medaṃ</i> ³
	-e + e- > -e-	51 <i>me + etaṃ</i> > <i>metaṃ</i>
-o)	-o + a- > -o-	18 <i>duddhakhīro + ahaṃ</i> > <i>-khīrohaṃ</i>
	-o + a- > -ā-	378 <i>-cchaddo + asi</i> > <i>-cchaddāsī</i> ⁴
	-o + a- > -vā-	p. 21,16 <i>kho + ahaṃ</i> > <i>khvāhaṃ</i> ⁵
	-o + aCC- > -aCC-	61 <i>appo + assādo</i> > <i>appassādo</i> ⁶
	-o + aCC- > -oCC-	25 <i>bhatako + asmi</i> > <i>bhatakosmi</i>
	-o + aCC- > -vaCC-	819 <i>kho + assa</i> > <i>khvassa</i>
	-o + aCC- > -vāCC-	998 <i>so + ajja</i> > <i>svājja</i>
	-o + aNC- > -aNC-	1040 1042 <i>ubho + antaṃ</i> > <i>ubhantaṃ</i> ⁷

¹ See Sv 389,10.

² cf. 1046 *kāme abhijappanti* > *kāmābhijappanti*; 1116 *dve + ahaṃ* > *dvāhaṃ*; and Ja V 154,18* *ve + ahaṃ* > *vāhaṃ*.

³ cf. p. 21,16 *dhamme + iti* > *dhammeti*. This is, however, probably not the result of -e + i-, but is an analogical form based upon the development of -i + iti. See the note on -i + i- above.

⁴ cf. 461 *rato + ahaṃ* > *ratāhaṃ*; 690 *anuttaro + ayaṃ* > *anuttarāyaṃ*; 692 *orako + ayaṃ* > *orakāyaṃ*; and S I 238,22 *kuddho + ahaṃ* > *kuddhāhaṃ*; Dh p 56 *yo + ayaṃ* > *yāyaṃ*; S II 178,8 *anamataggo + ayaṃ* > *anamataggāyaṃ*.

⁵ cf. Vin II 181,6 *ko + ayaṃ* > *kvāyaṃ*.

⁶ cf. 94 *asanto + assa* > *asantassa*; 324 *sammāniviṭṭho + assa* > *-niviṭṭhassa*; 584 *bhiyyo + assa* > *bhiyyassa*; 920 *anejo + assa* > *anejassa*; 922 *lolo + assa* > *lolassa*; 923 *phuṭṭho + assa* > *phuṭṭhassa*; 925 *pādalolo + assa* > *pādalolassa*; 994 *soko + assa* > *sokassa*; and Dh p 389 *yo + assa* > *yassa*.

⁷ cf. S III 93,9* *otiṇṇo + amhi* > *otiṇṇamhi*.

-o + ā- > -ā-	1136 <i>tamanudo</i> + <i>āsino</i> > - <i>nudāsino</i>
-o + i- > -i-	Ja IV 480,10* <i>mokkho</i> + <i>ito</i> > <i>mokkhito</i>
-o + i- > -o-	87 <i>yo</i> + <i>idha</i> > <i>yodha</i> ¹
-o + i- > -uvi-	J V 237,14* <i>ko</i> + <i>idha</i> > <i>kuvidha</i> ²
-o + eCC- > -eCC-	448 <i>vāyaso</i> + <i>etto</i> > <i>vāyasetto</i>
-o + o- > -o-	214 <i>yo</i> + <i>ogahane</i> > * <i>yogahane</i> ³

Excursus on the sandhi of *api* and (*i*)*ti*

In 460 *api* + *eva* > *apy eva* > *app eva*, i.e. the Pāli sandhi is a development of the Skt sandhi, showing the assimilation of *py* > *pp*.

In 241 *iti* + *eva* > *ity eva* > *icceva*; 251 *iti etam* > *iccetam*; 355 *iti* + *abravī* > *iccabravī*, i.e. the Pāli sandhi is a development of the Skt sandhi, showing the palatalisation of *ty* > *cc*.

We also find some contexts where what must originally have been (*i*)*ti* occurs as *tv*, as though from *tu*:

kālakato tv eva saṅkham gacchati, M III 137,8 foll.

Ekadhammasavanīyo tv eva samaññā ahoṣi, Th-a I 162,12.

As the first of these comes only a few pages after *ti* + *eva* > *t' eva*, it seems suspicious. Trenckner notes⁴ that “in canonical writings there is sufficient authority to distinguish between *t' eva* = *ti eva* and *tveva* = *tu eva*. In comments the latter is never used, but the scribes often substitute

¹ cf. 142 *brāhmaṇo* + *ti* > *brāhmaṇoti*. This is, however, probably not the result of -o + i-, but is an analogical form based upon the development of -i + *iti*. See the note on -i + i- above.

² If genuine, this is presumably a secondary development from *kv* with a svarabhakti vowel. Perhaps we should read *kvidha* with Be.

³ See *GD* II, note ad 214.

⁴ At Mil 423 (ad Mil 114,11).

tveva for *t' eva*. I do not scruple to correct it, though all my Mss give the wrong form throughout”.

I suspect that the reason for this is a development of *ti* > *tti*, as in AMg, which is elided to *tt'* before a vowel and is then “backformed” > *tv* by someone who had some knowledge of Sanskrit.

We also find *itveva* for *iti eva*, i.e. it is taken as though from *ito eva*¹:

itveva coro asim āvudhañ ca, M II 100,5* (*itvevā ti evam vatvā yeva*, Ps III 333,23, with v.l. *iccheva* [sic]) = Th 869 (*icceva* [Ee so] *iti eva, evam vatvā anantaram eva*, Th-a III 58,26).

Shepreth

K.R. Norman

¹ See CPD s.vv. *itveva* and *itvevam*, where it is derived from *ity eva(m)*, and the equivalence is *icceva*. See also Sadd 617,14–17.

APPENDIX

-a-	<	-aṃ + a- -i + a-	-āCC-	<	-a + aCC- -aṃ + aCC- -ā + aCC-
-aCC-	<	-a + aCC- -aṃ + aCC- -i + aCC- -e + aCC- -o + aCC-	-i-	<	-a + i- -i + i- -o + i-
-aNC-	<	-a + aNC- -aṃ + aNC- -ā + aNC- -e + aNC- -o + aNC-	-iCC-	<	-a + iCC- -ā + iCC- -i + aCC- -iṃ + aCC-
			-iNC-	<	-a + iNC- -ā + iNC-
-ā-	<	-a + a- -a + ā- -aṃ + a- -aṃ + ā- -ā + a- -ā + ā- -i + a- -i + ā- -iṃ + a- -u + a- -u + ā- -uṃ + a- -e + a- -o + a- -o + ā-	-ī-	<	-a + i- -aṃ + i- -ā + i- -i + i- -iṃ + i- -u + i-
			-u-	<	-a + u- -aṃ + u-
			-uCC-	<	-a + uCC- -aṃ + uCC-
			-uvi-	<	-o + i-

-ū-	<	-a + u- -aṃ + u- -u + i-	-vā-	<	-u + ā- -o + a-
			-vāCC-	<	-o + aCC-
-e-	<	-a + i- -a + e- -ā + e- -i + e- -e + a- -e + i- -e + e-	-ve-	<	-u + e-
-eCC-		-o + eCC-			
-o-	<	-a + u- -a + o- -ā + o- -o + a- -o + i- -o + o-			
-oCC-	<	-o + aCC-			
-yaCC-	<	-e + aCC-			
-yāCC-	<	-e + aCC-			
-va-	<	-u + a-			
-vaCC-	<	-o + aCC-			